Do Futures Prices Help Forecast the Spot Price?

Xin Jin*

Department of Economics, Business School, University of Aberdeen

March 15, 2017

Abstract This paper proposes a futures-based unobserved components model for commodity spot prices. Prices quoted at the same time incorporate the same information, but are affected differently, resulting in the different shapes of futures curves. This model utilizes information from part of the futures curve to improve forecasting accuracy of the spot price. Applying this model to oil market data, I find that the model forecasts outperform the literature benchmark (the no-change forecast) and futures prices forecasts in multiple dimensions, with smaller average error variation over the sample period and higher chance of smaller absolute error in each period.

Keywords: Forecasting; Commodities; Spot Price; Futures Price; Futures Curve; Unobserved Components Model; Stochastic Process

JEL Classification Numbers: G17, Q4, G13, C38.

*Address: Edward Wright Building S67, Dunbar Street, Old Aberdeen AB243QY UK, e-mail: *xjin@abdn.ac.uk*. The author is grateful to the editor, the anonymous referee, Nathan Balke, Tom Fomby, Mine Yücel, Marc Gronwald and Euan Phimister for advice and suggestions. All errors and mistakes are the author's.

1 Introduction

Given the high volatility of commodity prices and the importance of raw materials in production, accurate forecast of the spot price is of great interest for various purposes. Policy makers and central banks closely track commodity prices, especially crude oil prices. Price forecasting is also crucial to business decisions in many industries.

One intuitive forecast of the spot price is the futures price. The efficient market hypothesis suggests the futures price as the best forecast. A large literature has discussed the forecasting ability of futures prices. French (1986), Fama and French (1987), Bowman and Husain (2004), Coppola (2008), Reichsfeld and Roache (2009), Reeve and Vigfusson (2011), Chinn and Coibion (2014) among others find evidence confirming the forecasting ability of the futures prices for certain commodities, while equally large amount of research like Bopp and Lady (1991), Moosa and Al-Loughani (1994), Chernenko et al. (2004), Alquist and Kilian (2010), Alquist et al. (2013) find little evidence supporting the futures price as the best forecast. Instead, no-change forecast is suggested as a plausible measure of the expected spot price. While Alquist et al. (2013) provide a comprehensive overview of various forecasting models including futures-based ones, they do point out the potential of factor (unobserved components) models for forecasting.

This paper contributes to the forecasting and price dynamics literature by proposing a futures-based unobserved components forecasting model which has superior forecasting accuracy. The improved forecasting accuracy relies on identifying the spot price stochastic process by exploiting part of the futures curve (the term structure of the futures prices). The shape of the futures curve is partly determined by the spot price stochastic process, if we consider the futures price as composed of the expected future spot price and risk premium as in Pindyck (2001).

Prices quoted at the same time for immediate and future delivery of the same commodity all incorporate the same set of information under the efficient market hypothesis, as argued by earlier work like Working (1942), Tomex and Gray (1970) and Peck (1985). However the prices are affected by the same set of information differently. For example, the information set includes the growing demand for a certain commodity from the emerging economies, and the exploring and drilling activities in search of it, which would have long-lasting effects on the prices, as well as temporary supply shortage and short-term demand increase, which would have short-lived effects on the prices. Thus the prices with further delivery dates are much less affected by the short-term effects of the information compared to the prices with nearer delivery dates, while all prices are affected by the long-term effects similarly. The difference between the prices with further delivery dates (the far end of the futures curve) and nearer delivery dates (the near end of the futures curve with the nearest being the spot price) roughly reflects the the short-term effects of the information, while futures prices with further delivery dates (the far end of the futures curve) roughly reflect the long-term effects of the information. Thus futures curve helps infer the stochastic process of the spot price.

The model allows for flexibility to fit the rich dynamics of the spot price and futures curves, while it is still relatively easy to estimate. Applying the model to oil market data, I show that the model forecasts outperform both the no-change and the futures price forecasts in multiple dimensions. Using 5-year rolling-window out-of-sample forecasting over 20 years at weekly frequency, the model outperforms the literature benchmark (the no-change forecast) and the futures prices forecasts especially at longer forecasting horizons from 32 to 48 weeks, as measured by MSE, ME and MAE. The improvement at longer forecasting horizons is also statistically significant as tested by the finite-sample unconditional predictive ability test developed in Giacomini and White (2006). In addition, the model performs better in the 2000s than in the 1990s. The model suggests this could be due to the changing commodity market conditions.

This paper differs from recent research that has already paid attention to the value of futures curves in the assumption of the spot price stochastic process. Coppola (2008) proposes a VECM model which essentially uses futures curves for forecasting the spot price, where the spot price is modeled as a random walk to which the effects of shocks will never dissipate. To the contrary, this paper argues the effects of shocks to the spot price could partially dissipate over time.

Motivated by the cost of carry model, the two-factor and three-factor models proposed in Schwartz (1997) and Cortazar and Schwartz (2003) also bear resemblance to my model as they are essentially estimated using futures curves. However, the models intuitively imply that the spot price follows a random walk process in the discrete time. The assumption implies the effect of shocks to the spot price will not dissipate over time. This is fundamentally different from the proposed model in this paper.

This paper instead assumes that the spot price dynamics contains some temporary component. The idea that shocks to the spot price could be dissipated is not new. Both theoretical and empirical works suggest that commodity price contains both a long-term component and a short-term component (see Fama and French (1988), Pindyck (1999), Carlson et al. (2007) and Stevens (2013)). Allowing part of the shocks to dissipate, the unobserved components model proposed in this paper can then be seen as the empirical extension of the recent development in the commodity price theory by Carlson et al. (2007) and Stevens (2013). In terms of the assumed the spot price stochastic process, my model is similar to the model of price evolution in Pindyck (1999), but further relates the futures prices of different maturity terms to the spot price, and thus is able to exploit the information from futures curves.

The plan of the paper is as follows. Section 2 develops the unobserved components model of crude oil prices and discusses the model features. Section 3 provides an overview of the data, describes the statistical tests adopted and presents the estimation results from oil market. Section 4 presents and evaluates the forecasting ability of the model by comparing it to alternative benchmarks and over time. Section 5 concludes.

2 Modelling of Commodity Prices

In this section, an unobserved components model is constructed. The unobserved components model is able to take full advantage of the information in the futures curve and uncover the unobserved factors affecting the prices.

2.1 The Spot and Futures Prices Stochastic Process

Shocks to commodity spot price could have effects of different time-persistence. French (1986) discusses the possibility that "shocks to the current price are dissipated before they can affect the expected price" and observes that in such scenario futures prices can provide "reliably better forecast". Such observation indicates that shocks to the spot price would dissipate over time, rather than being permanently preserved as in a martingale walk process. Fama and French (1988) discuss the long-term and short-term components in asset prices although the purpose there is to test market efficiency. Pindyck (1999) also argues that a model of price evolution should incorporate both a reversion to a long-run trend and a non-stationary stochastic long-run trend after studying both the empirical features of commodity prices and the theoretical implications of the depletable resource prices. More recently, Carlson et al. (2007) demonstrate that such price dynamics with both long-term and short-term components could arise from a hotelling model with production adjustment costs. The impulse response functions of the spot price to shocks in Carlson et al. (2007) show that part of the shocks are incorporated into price as temporary increments, rather than all shocks having long-lasting effects on the price. Stevens (2013) also derives similar price dynamics from a hotelling model with storage.

Thus this paper assumes the spot price to contain both long-term and short-term components. This is similar to Pindyck (1999), as both indicate partially dissipating shocks to the spot price over time while overall the price maintains non-stationarity.

As to the the spot and futures prices interaction, there have been two views. The first is based on the cost-of-carry model which views the expected spot price as being different from the spot price by interest rates, convenience yield and storage costs (see for example Brennan (1958)). This is also referred to as the theory of storage. In the empirical works based on this view, the futures price is either used as the proxy of expected spot price (see for example Fama and French (1987)), or linked directly to the spot price with a slightly differently defined convenience yield (see for example Pindyck (2001) and Figuerola-Ferretti and Gonzalo (2010)). The other view models the futures price as the sum of the expected spot price and a risk premium (see for example Pindyck (2001)). The two views are not exclusive of each other, and both involve the concept of the unobserved expected spot price.

Since this paper models explicitly the spot price dynamics, which implies the unobserved expected spot price, the futures price is modelled following the second view. The following section will discuss the modelling of the spot and futures prices based on the above assumptions.

2.2 An Unobserved Components Model of Spot and Futures Prices

Like in Pindyck (1999), the long-term component reflects the time-varying long-run equilibrium price level, and the short-term component reflects the short-term deviation from this long-run equilibrium level. Neither component is observable. The spot price generating process is summarized by the following equations:

$$p_t = \tau_t + c_t + \epsilon_t^p \qquad \epsilon_t^p \sim N(0, \sigma_p^2) \tag{1}$$

$$\tau_t = \tau_{t-1} + \epsilon_t^\tau \qquad \epsilon_t^\tau \sim N(0, \sigma_\tau^2) \tag{2}$$

$$c_t = \rho_1 c_{t-1} + \rho_2 c_{t-2} + \epsilon_t^c \qquad \epsilon_t^c \sim N(0, \sigma_c^2) \tag{3}$$

where p_t is the spot price at time t, τ_t is the non-stationary long-term component, c_t is the stationary short-term component, and ϵ_t^p is an idiosyncratic noise term in the spot price. ϵ_t^τ and ϵ_t^c are idiosyncratic noise terms in the long-term and short-term components respectively, and are assumed to be correlated with coefficient $cov_{\tau,c}$.

For the futures market, futures prices do not contain more information about the future compared to the spot price (see for example Working (1942), Tomex and Gray (1970) and Peck (1985)). The information incorporated in futures prices is the same as in the spot price. Thus the T-period futures price can be written as:

$$f_t^T = E_t\Big(p_{t+T}\Big) + RP_t^T + \epsilon_t^{f^T} \qquad \epsilon_t^{f^T} \sim N(0, \sigma_{f^T}^2) \tag{4}$$

where f_t^T is the price at time t for a contract maturing at $t+T$, RP_t^T is a term incorporating all the risk at time t associated with the futures contract of a maturity term T, and $\epsilon_t^{f^T}$ is an idiosyncratic noise term in the futures price. Neither $E_t(p_{t+T})$ or RP_t^T is observable. However $E_t(p_{t+T})$ can be inferred if the spot price stochastic process is known[1].

I model the risk premium term RP_t^T as follows:

[1] Details are provided in Appendix A

$$RP_t^T = \mu_{rp}^T + \beta_T rp_t \tag{5}$$

$$rp_t = \rho_{rp} rp_{t-1} + \epsilon_t^{rp} \qquad \epsilon_t^{rp} \sim N(0, \sigma_{rp}^2) \tag{6}$$

where μ_{rp}^T is a constant capturing the term-dependent average risk premium level for a futures contract with a maturity term T, which is time-invariant; rp_t is an AR(1) component capturing the time-varying part in the risk premium at the price-quoting time t, and is common to all futures contracts, regardless of different maturities. However the common risk premium component rp_t is loaded into each futures price's risk premium term with term-dependent loading factor β_T. This allows for handling futures prices with different maturity terms, and also allows for term-specific time-varying risk premiums.

The equation for the price of a futures contract with a maturity term T follows through naturally:

$$f_t^T = E_t\left(p_{t+T}\right) + \mu_{rp}^T + \beta_T rp_t + \epsilon_t^{f^T} \qquad \epsilon_t^{f^T} \sim N(0, \sigma_{f^T}^2) \tag{7}$$

$$rp_t = \rho_{rp} rp_{t-1} + \epsilon_t^{rp} \qquad \epsilon_t^{rp} \sim N(0, \sigma_{rp}^2) \tag{8}$$

The two sets of equations for the spot and futures prices hold simultaneously in an efficient market. Equations (1), (2), (3), (7) and (8) should be estimated as a system with observed p_t and f_t^T. The model then can be then rewritten into a state-space form and estimated by maximum likelihood using the Kalman filter. The details of the state-space form and Kalman filter are provided in Appendix A. Before moving on to estimation, I will first discuss the spot and futures price dynamics implied by this model.

2.3 Model Discussion

The model provides rich dynamics to fit most features of the observed price dynamics like in Figure 1 from crude oil market. One implication from the model is that it implies lower volatility for the prices of future contracts with longer maturity terms. Under the model, all prices contain a non-stationary long-term component as well as a stationary short-term component, but are affected by them differently. In other words, the spot and futures prices all contain the same set of information but are affected differently. Futures prices, especially those with longer maturity terms, are less affected by the short-term shocks like temporary supply disruptions. The longer the maturity term, the more the short-term shocks would dissipate, leaving the futures prices to be affected by mostly the long-term shocks. This implication is consistent with observed market data.

Another implication from the model is that it allows for flexible futures curves. Futures curves simulated by the model can be upward or downward sloping, or even hump-shaped. If current spot price is lower than the long-term level, the negative short-term component's effects on future delivery would be expected to dissipate over time, resulting in higher futures price quoted today, and thus a upward-sloping futures curve. The futures curve could also be further shaped by the futures contract risk premiums. If the risk premium for futures contracts with longer maturity terms is more negative than the shorter ones, the different risk premiums for different maturity terms might result in a hump-shaped futures curve.

Furthermore, the model provides an alternative interpretation of forecasting accuracy: the market conditions affect the forecasting accuracy. If the spot price dynamics is indeed more than a random walk[2], the existence of stationary component would reduce the model forecasting error variation relatively to a no-change random walk forecast with certainty, especially at longer forecasting horizons. More specifically, the reduction in the relative MSE of the two is determined by the volatility of the short-term component.

To see this, the T-step ahead forecast at time t from the model $F^{t,T}_{model}$ can be written as the function the long-term and short-term components in the following way:

$$F^{t,T}_{model} = E_t\left(p_{t+T}\right) = \tau_t + E_t\left(c_{t+T}\right) \tag{9}$$

while the realization of the actual spot price at $t + T$ would be affected by the volatilities of the long-term and short-term components. As a result, the forecasting accuracy could be largely affected by changing market conditions like the prevalence of short-term or long-term shocks and their sizes in the market.

On the other hand, the random walk forecast $F^{t,T}_{rw}$ and the futures price forecast $F^{t,T}_{f}$ are:

[2] It is noteworthy that the finance literature argues the speculative prices follow a more general martingale process, of which the random walk is a special case.

$$F_{rw}^{t,T} = p_t = \tau_t + c_t + \epsilon_t^p \tag{10}$$

$$F_f^{t,T} = f_t^T = E_t\left(p_{t+T}\right) + RP_t^T + \epsilon_t^{fT} \tag{11}$$

Comparing Equation 9 and Equation 10, the forecast difference $(E_t(c_{t+T}) - c_t - \epsilon_t^p)$ comes in part from the dissipated short-term component $(E_t(c_{t+T}) - c_t)$. The larger the T, the greater the difference, as $E_t(c_{t+T})$ converges to zero.

3 Empirical Studies: the Case of Crude Oil Price

3.1 Model Estimation

Technically the model can be estimated with a flexible number of different prices. In the estimation algorithm, the price equations are written in matrix, with the same rows in all matrices corresponding one price equation. The dimension of the matrices can be adjusted depending the number of different prices used in the estimation. Appendix A illustrates the matrix form when one futures price and the spot price is used.

Among the different futures prices, the long-term futures prices on the futures curve provides information of the long-term component, since the model implies that the long-term futures prices contain mostly the long-term component (τ_t) and their risk premiums. On the other hand, the time-varying differences among different prices provide information on the dynamics of the short-term and risk premium components (c_t and rp_t). Together they determine the shape of the futures curves, and make up the "term structure" that provide information to identify the unobserved components.

3.2 Data Overview

I use the crude oil data for applying the model. The main advantage of crude oil market is the availability of data. WTI is traded on both the spot and futures markets. The crude oil producers and refineries use the futures market for hedging risks associated with the production, processing and handling of crude oil. Presumably all WTI prices would be affected by the same market disturbances, which is assumed by the proposed model.

Furthermore, although most futures trading volume is concentrated in nearby contract months in many futures markets in the West, some futures markets like crude oil and natural gas go out up to twelve years into the future. Specifically, the WTI futures contracts have various maturity terms, from 1-month maturity to 78-month maturity. Figure 1 plots the crude oil spot price along with futures curves (up to 35-month maturity) from 1989 to 2015.

Even so, trading volume and liquidity falls sharply after the first 6 to nine months with many distant deferred futures contract months experiencing zero trading volume on any given day. Exchanges will often post "indicative" settlement prices on days when no trading occurs. Economic theory suggests that only market transaction prices contain relevant information. The relative trading volume across contract months for 3 typical days is shown in Figure 2.

As a result, the prices selection should balance between long enough maturity lengths and enough actual non-zero trading, and roughly reflect the shape of the curve. Anecdotal evidence from the industry shows that some oil producers hedge for 2 years (24 months) ahead, indicating that the 24-month (or longer-maturity) futures price, if it is a price from actual trading, contains information of the long-term component in the price.

However, the limited trading of longer-maturity futures contracts limits the longest maturity length to be used, as the "indicative" prices reported do not contain quality information of the long-term component. For example, the 24-month futures contracts whose trading started in September 1995, had zero trading for more than 30% of the weeks until November 2014.

In this application the model is estimated with WTI spot and 6-, 12-, 18-months futures prices at weekly frequency[3] from the week ending on March 31, 1989 to the week ending on November 28, 2014[4]. The prices are available from Energy Information Administration website and Datastream. The 6-month, 12-month and 18-month futures prices used in the estimation have their trading volume plotted in log in Figure 3.

Overall 18-month futures price could be a balance between a long enough maturity length for uncovering the long-term component of oil prices and a large size of a quality sample. Herce et al. (2006) argue that a 18-month length may be far enough to capture the long-term component as most of the short-run fluctuations in

[3]The prices are the daily closing prices at the end of each week.

[4]Using different data (WTI spot and 3-, 6-, 12-month futures prices) leaves the forecasting results qualitatively unchanged. Shorter sample periods are also used and the results are robust.

prices dissipate within this horizon. And this price series contains more than 90% of observations from actual trading.

As discussed earlier, the crude oil price data contain market data, indicative data and missing observations. Market data are the prices from actual trading (reported and traded), indicative data are the reported price when no trading occurs (reported but not traded)[5], and missing observations are the "NA's" (not reported and not traded).

Throughout the sample period of 1340 weeks in total, there are 129 weeks of no trading, all in the 18-month futures contracts. In terms of the price data, there are 129 missing or indicative observations in total. Specifically, the 18-month futures price series has 49 periods of no observations from March 31 1989 until 1995 (i.e. 25 consecutive "NA's" for the beginning, then another occasional 24 "NA's"). From August 11 1995 until November 28 2014 (1008 weeks), it contains 80 indicative observations (reported but not traded). One advantage of the model is its flexibility with missing observations in data. More details of the algorithm are provided in Appendix A.

Weekly frequency is selected since intuitively a lot of very-short-run disturbances to the price would be averaged out at monthly frequency. Also, the crude oil market fundamentals like supply, demand and inventory drives the spot price dynamics. Their data are published at monthly and weekly frequency, and would be reflected in the weekly data. Comparing the spot and futures prices at weekly frequency

[5]The phenomenon of indicative prices is common across futures exchanges whenever no futures transactions occur. In the case of WTI futures contracts, the New York Mercantile Exchange provids the indicative prices to Datastream.

could help separate the short-term component in the price, which could potentially improve the forecasting accuracy.

Although the model could work for both real and nominal prices, intuitively it seems more appropriate to measure the prices in real term, as the long-term component for the equilibrium price ultimately is determined by the market fundamentals like oil demand and supply. US CPI data is available from U.S. Bureau of Labor Statistics at monthly frequency, which is linearly interpolated to weekly frequency.

Persistence test results using ADF test are reported in Table 1. As the results show, all data series contain a unit root, confirming the view of the model that the spot and futures prices contain a long-term random-walk component.

3.3 Estimation Results

In this section, I estimate the model using the full sample from 1989 to 2014 with time-varying risk premiums and correlated long-term and short-term components ((12) and (15) in Appendix A.). Estimates are reported in Table 2.

The point estimates of most parameters are significant at 99% significance level. Due to the small magnitude of the point estimates, I also plot the estimated unobserved components time series with 90% confidence intervals in Figure 4: long-term τ_t, short-term c_t, and the time-varying component of risk premiums rp_t. I also plot overall risk premiums for different contracts in Figure 5. Note that the estimated risk premiums from all contracts are systematically negative throughout the sample period, and tend to be less negative post 2000 except for a short period around 2007 - 2008 and towards the end of the sample period.

Keynes (1930) theory of risk premium proposes that if hedgers in the futures market are net short (i.e. producers of the physical commodity) seeking price protection, then in order to entice speculators into taking the offsetting long positions, the risk premium (as define as in the model) would be negative as a reward. The longer the maturity term, naturally the higher the risk associated with the futures contract, thus the reward for bearing the risk would tend be larger in size. When the hedgers are net long (i.e. buyers), risk premium would be positive. It would follow from the theory that the sign and size of risk premium would logically be related to the distribution of hedgers.

This model enables estimating risk premiums by exploiting the spot price and futures curves, rather than approximating the unobserved risk premiums as the naive difference between the spot and futures prices. The resulting estimated risk premiums share similar pattern as in recent literature. Even without splitting the sample period into two as in Hamilton and Wu (2014), the estimated risk premiums plotted in Figure 5 show similarly smaller on average compensation to the long position in more recent data (the risk premiums become less negative). In my results, however the very end of the sample period sees a larger compensation to the long position again.

4 Out-of-Sample Forecasting Performance

4.1 Overall Forecasting Accuracy of the Model

The out-of-sample forecasting accuracy is mainly evaluated in this paper by four measures, and the statistical significance of the improvement is also tested. MSE measures the average variation of the forecasting errors, ME measures the overall unbiasedness of the forecasts, MAE provides an overview of the average absolute error size. Although the first three measures are widely used in the literature and provide a good overview, a relatively low (high) MSE can be driven by a minority of extremely good (bad) forecasts. Thus in addition I also provide the fourth relative measure, that is the fraction of relatively smaller absolute errors. The benchmark alternatives are the no-change and the futures prices forecasts.

Detailed summarizing statistics of the forecasting errors in Table 3 reveal that overall the model improves the forecasting accuracy, especially at forecasting horizons longer than 12 weeks. Using the forecasting unbiasedness measure ME, the model outperforms the two alternatives at all forecasting horizons. Using the average forecasting error variation and size measures MSE and MAE, the model outperforms the two alternatives at all forecasting horizons longer than 12 and 8 weeks, respectively.

Among the two alternative estimation methods, the forward recursive estimation is better using MSE at all forecasting horizons longer than 8 weeks, while the rolling-window estimation generates better forecasts using ME and MAE at all forecasting horizons longer than 4 weeks.

To test statistical significance of the improvement, I carry out the finite-sample unconditional predictive ability test (Giacomini and White (2006)) and compare the forecasts from the rolling-window estimation and the two alternative benchmarks[6]. Column 1 in Table 3a-c shows that, comparing the model and the random walk forecasts by MSE, the null hypothesis of equal predictive accuracy of the two is rejected at 10% significance level at the 48-week forecasting horizon, and the model outperforms the random walk with 8% lower MSE; by MAE the null is rejected at 10% to 1% significance levels at multiple forecasting horizons (24- to 48-week), and the model outperforms the random walk with 5% to 10% lower MAE. Comparing the model and the futures prices forecasts by ME, the null hypothesis is rejected at 10% significance level at multiple forecasting horizons (32- to 48-week) and the model outperforms the futures prices with 40% closer-to-zero ME; by MAE the null is rejected at 10% significance level at the 28-week forecasting horizon and the model outperforms the futures prices with 4% lower MAE. On the other hand, except for the 4-week forecasting horizon comparison between the model and the futures price by MSE and MAE, in no other cases has evidence of statistically better forecasts by either the random walk or the futures prices been found. Figure 6 illustrates the actual spot price with the model and the alternative forecasts and shows that the model forecasts are closer to the actual price.

In addition I also present the fraction of model improvement in Table 3d. The fraction of the model forecasts with smaller absolute errors relative to the random walk and the futures prices alternatives adds another dimension of forecasting ac-

[6] A comparison between the forecasts from the forward recursive estimation and the two alternatives is not allowed using Giacomini and White (2006).

curacy. Overall the model forecasts have smaller absolute errors relative to both alternatives more often at longer forecasting horizons. The forward recursive estimation performs better than the rolling estimation when measured relative to the futures prices, while the rolling estimation performs better when measured relative to the random walk.

The forecasting evaluation in Table 3 shows strong evidence of superior performance of the model[7]. At longer (than 12 weeks) forecasting horizons, the model forecasts have higher chance of smaller absolute errors, have average smaller absolute errors and variation, and are more unbiased.

Table 4 reports the robustness of the results. The model is estimated with different price data and compared to the literature benchmarks: one with the spot and the 3-, 4- and 6-month futures prices, one with the spot and the 3-, 6- and 12-month futures prices. Similar to the results in Table 3, the model forecasts outperform the benchmark forecasts especially at longer forecasting horizons, and the performance improvement is statistically significant when 12-month and/or 18-month prices are used.

4.2 How the Model Improves the Forecasts

The proposed unobserved components model argues that the futures curve and its changes over time provide useful information about price movements. Under the assumption of no-arbitrage, futures curves provide important additional information about the evolvement of the spot price than the spot and futures prices individually.

[7]Using different subsamples (1989-1995, 1995-2014, 2008-2014) leaves the results qualitatively unchanged. The subsamples reflect different stages of futures trading volume.

The addition of such information improves the forecasting accuracy relative to both the random walk and the futures prices forecasts.

Furthermore, such improvement relies on more than just a simple composite of different futures prices. For example, in Table 3b, using ME, at 36-week and 48-week forecasting horizons the futures prices perform worse than the random walk. Nonetheless at both forecasting horizons the model forecasts outperform both the random walk and the futures prices forecasts. Also, the improvement in forecasting accuracy is not only in terms of smaller average absolute error size and variation, but also in terms of the higher chance that the model forecasts are better.

However so far the long-term and short-term view of spot price movements and the additional information from futures curves have been largely ignored in the literature of forecasting and modelling commodity price dynamics. The proposed unobserved components model provides a way to utilize futures curves and proves to generate more accurate forecasts compared to the literature benchmarks.

4.3 Time-varying Forecasting Accuracy of the Model

Although the overall forecasting accuracy improvement is convincing, one interesting question is whether the predictability has changed over time.

To have a closer look, I choose the first 1032 forecasts at different forecasting horizons (starting from March 18, 1994 to December 20, 2013) and calculate MSE and ME every 12 weeks. Figures 7 and 8 plot the approximately 3-month average forecasting accuracy over almost 20 years (86 MSEs and MEs for each forecasting horizon). In these figures, the average forecasting accuracy of the model is relatively consistent over time except for mid-2008. At longer forecasting horizons, average forecasting errors during the 2000s are slightly smaller compared to the 1990s.

However, smaller MSE and ME do not necessarily mean that the forecasting accuracy is improved the 2000s. It could be by chance that the forecasting errors tend to be slightly smaller on average in the 2000s, regardless of what forecast is used. A comparison between the model and the random walk forecasts is presented in Figure 9 and 10, which plot out the MSE and ME ratios of the model forecasts to the random walk benchmark. A ratio smaller than 1 in absolute terms indicates an improvement of the model.

Figure 9 and 10 show that in the 2000s the model forecasts more often have smaller MSE and ME compared to the random walk. During mid-2008 when forecasting errors are large for all forecasting measures, the model forecasts are still comparable to the random walk counterparts.

The comparison shows the model's relative performance is slightly better in the 2000s than earlier. This observation is very different from Chinn and Coibion (2014) who document a broad decline in the predictive content of commodity futures prices since the early 2000s. More importantly, my model provides a very different interpretation of the time-varying forecasting accuracy. The model implies that the reduction of the model forecasting error variation relative to a random walk depends on the short-term component volatility. The higher the short-term component volatility, the lower the MSE and ME ratios of the model forecasts over the random walk, and the more the model outperforms the random walk, especially at longer forecasting horizons. The estimated short-term volatility σ_c^2 from the 260-week (approximately

5-year) rolling-window estimation shows that the short-term volatility volatility during the 2000s is indeed larger than before, except for the end of sample period [8]. Both this observation and the relative forecasting accuracy improvement in the 2000s are consistent with the model's prediction.

In other words, the better relative forecasting performance could be simply driven by the changing market conditions. For example, long periods of strong demand with constrained supply in the 2000s might result in more temporary shortage of crude oil, and thus higher volatility of the short-term stationary component, which reduces the relative forecasting error variation of the model.

The changing market conditions have also been documented and discussed in the price discovery literature (for example Silvério and Szklo (2012) and Caporale et al. (2014)), but this model provides an alternative interpretation as the time-varying price dynamics is driven by the short-term component in the spot price, which might be driven by commodity market fundamentals like supply and demand, rather than by the futures market efficiency.

5 Conclusion

This paper proposes a model of commodity spot and futures prices for forecasting. The improved forecasting accuracy from the model is convincing. When applying the commodity price model to 25 years of oil market data, the model forecasts outperform the literature benchmarks in multiple dimensions. During the sample period, the model forecasts are more unbiased on average, have smaller forecasting error size

[8] Detailed estimates can be provided on request.

and variation on average, and are more often so when compared to both the random walk and the futures prices forecasts.

The superior performance is a result of utilizing the additional information from the futures curves and the careful approximation of the expected spot price. The spot and futures prices quoted on the same date contain the same information, but are affected to a different extent. More specifically, there is information with longer lasting effects on the prices, and also information with temporary effects. I capture this difference by decomposing the spot price into a non-stationary "long-term" component and a stationary "short-term" component. This spot price dynamics model enables a more careful approximation of the expected future spot price, which is the model forecast. Thus the model forecast is able to outperform the random walk forecast, which assumes that the spot price only contains the long-term component and has the same forecast for all forecasting horizons. As the forecasting exercises show, the advantage of the model forecasts is indeed more apparent at longer forecasting horizons, when the short-term effects dissipate.

Furthermore, the model allows for estimating risk premiums in futures prices. Thus the model forecast also outperforms the naive futures price forecast, which contains both the expected future spot price and a risk premium. As the results demonstrate, the model forecasts outperform both the random walk and the futures prices forecasts.

The proposed model challenges the notion that the forecasting ability of the commodity futures market is an indicator of the market efficiency (see for example Malkiel (2003) and Chinn and Coibion (2014)). The time-varying relative forecasting accuracy could merely reflect the changing short-term component volatility, in other words, the changing commodity market conditions, rather than the efficiency level of the futures market.

The model provides an approximation of the unobserved expected spot price, which is crucial for understanding the interaction between the futures and spot markets. The expected spot price is the cornerstone of concepts such as futures risk premiums and convenience yields (as defined in for example Brennan (1958)). A more accurately approximated expected spot price allows for a more accurate estimation of risk premiums in futures prices. While this paper does not address directly the fundamentals behind the changing market conditions and the micro-foundation of the structure of risk premiums, the resulting estimated risk premiums from the model share similarities with that of Hamilton and Wu (2014). Similarly this model could be used to infer the convenience yield, which is closely related to market fundamentals of storable commodities. Application of this model would prove useful for future work in the literature on futures risk premiums as well as convenience yields. Application of this model to other commodity data and data at daily frequency could also be explored in future research.

A State-space Setting and Maximum Likelihood Estimation using the Kalman Filter

This section shows that the proposed unobserved components model can be rewritten into a state-space form using the example of modeling the spot and one futures prices. It can be extended to include multiple futures prices with different maturity lengths as in the paper. It also provides more details about the maximum likelihood estimation with Kalman filter at the end.

The long-term component τ_t, short-term component c_t described by Equations (2) and (3), and the time-varying risk premium component rp_t serve as the unobserved states underlying the price dynamics. The state equation can be rewritten in matrix form as the following:

$$\begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \\ rp_t \end{bmatrix} = \mathbf{F} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ c_{t-2} \\ rp_{t-1} \end{bmatrix} + \mathbf{G} * 1 + \mathbf{Z}\nu_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \rho_1 & \rho_2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \rho_{rp} \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ c_{t-2} \\ rp_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \epsilon_t^{\tau} \\ \epsilon_t^{c} \\ 0 \\ \epsilon_t^{rp} \end{bmatrix} \quad (12)$$

The spot price equation can be rewritten as:

$$p_t = \tau_t + c_t + \epsilon_t^p = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \\ rp_t \end{bmatrix} + \epsilon_t^p \quad (13)$$

The futures price (with maturity term T) equation can be rewritten as:

$$
\begin{aligned}
f_t^T &= E_t\left(p_{t+T}\right) + \mu_{rp}^T + \beta_T rp_t + \epsilon_t^{f^T} \\
&= \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} (\mathbf{F})^T \begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \\ rp_t \end{bmatrix} + \mu_{rp}^T + \beta_T rp_t + \epsilon_t^{f^T} \\
&= (\begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} (\mathbf{F})^T + \begin{bmatrix} 0 & 0 & 0 & \beta_T \end{bmatrix}) \begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \\ rp_t \end{bmatrix} + \mu_{rp}^T + \epsilon_t^{f^T}
\end{aligned} \tag{14}
$$

where $\mathbf{F}$ matrix is the loading matrix governing state variables dynamics in Equation (12).

Equations (13) and (14) can be then fit into the observation equation as follows:

$$\begin{aligned}\begin{bmatrix} f_t^T \\ p_t \end{bmatrix} &= \mathbf{H} \begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \\ rp_t \end{bmatrix} + \mathbf{A} * 1 + w_t \\ &= \begin{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} (\mathbf{F})^T + \begin{bmatrix} 0 & 0 & 0 & \beta_T \end{bmatrix} \\ 1 \quad 1 \quad 0 \quad 0 \end{bmatrix} \begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \\ rp_t \end{bmatrix} + \begin{bmatrix} \mu_{rp}^T \\ 0 \end{bmatrix} + \begin{bmatrix} \epsilon_t^{f^T} \\ \epsilon_t^p \end{bmatrix}\end{aligned} \tag{15}$$

The state-space model is estimated using maximum likelihood with Kalman filter as follows (see for example Hamilton (1994), Durbin and Koopman (2012), Harvey (1989)): for a model of time series given by the following state equation for z_{t+1} and observation equation for y_t,

$$z_{t+1} = F z_t + u_t, \text{and } y_t = H' z_t + v_t \tag{16}$$

with serially uncorrelated $u_t \sim N(0, Q)$, $v_t \sim N(0, S)$, the state vector z_k at time k estimated by Kalman filter, given observations of y_t up to k, is represented by the posteriori state estimate $\hat{z}_{k|k}$ and the posteriori error covariance matrix $P_{k|k}$. $\hat{z}_{k|k}$ and $P_{k|k}$ are estimated by the recursive steps below ($\hat{z}_{k|k-1}$ represents a forecast of z_k based on no observations of y_k):

$$\hat{z}_{k|k-1} = F\hat{z}_{k-1|k-1} \tag{17}$$

$$P_{k|k-1} = FP_{k-1|k-1}F' + Q \tag{18}$$

$$\hat{z}_{k|k} = \hat{z}_{k|k-1} + P_{k|k-1}H(H'P_{k|k-1}H + S)^{-1}(y_k - H'\hat{z}_{k|k-1}) \tag{19}$$

$$P_{k|k} = P_{k|k-1} - P_{k|k-1}H(H'P_{k|k-1}H + S)^{-1}P_{k|k-1} \tag{20}$$

Equation 19 is referred to as the *updating equation*, and its coefficient matrix $P_{k|k-1}H(H'P_{k|k-1}H + S)^{-1}$ is known as the *Kalman gain*, which weighs observation y_k to update $\hat{z}_{k|k-1}$.

In addition to filtering the state vector, the Kalman filter also gives the likelihood function given the parameters. This way, the parameters can be estimated by maximum likelihood, starting from an initial guess of the parameters and the initial states.

The initial states for Kalman filter is set to be uninformative: the initial long-term component $\tau_{0|0}$ is set to be the average of the spot price[9]; others are set to be zeros. The covariance matrix for the initial states $P_{0|0}$ is set to be symmetric with 100 on the diagonal and zeros off the diagonal. Different initial guesses of the parameters have been tried; the results reported are based on the one with highest likelihood when convergence occurs[10].

[9] For rolling and recursive estimation for forecasting exercises, the average is of the specific spot price sample used for estimation

[10] For rolling and recursive estimation for forecasting exercises, the estimations are made iteratively and the initial guesses of the parameters are set to be the estimated parameters in the earlier iteration to speed up the estimation.

In the estimation using the spot and 6-, 12- ,18-month futures prices, the 18-month futures price is not always available. For these periods with "NA" observations, the Kalman filter algorithm is set up such that in the *updating equation* step (Equation 19), the observation equation matrices (H and S) are resized depending on what data are available and only contain the elements corresponding to the observed data. Intuitively this means in those periods, the Kalman gain puts no weight on the missing observation of the 18-month futures price and uses only the observed spot, 6-month and 12-month futures prices when updating the state vector estimate[11].

[11] Durbin and Koopman (2012) discuss missing observations in Section 4.10. See also Harvey and Pierse (1984) and the Dynamic Stochastic General Equilibrium literature like Giannone et al. (2008) where macroeconomic data used in the estimation are of mixed frequencies.

Table 1: **Overview of Crude Oil Prices Persistence**

(a) level

Series	ADF t-statistic	# of lags	ADF t-statistic	# of lags	ADF t-statistic	# of lags
WTI spot	-0.703	1	-0.705	2	-0.799	3
6-month	-0.339	1	-0.393	2	-0.550	3
12-month	-0.120	1	-0.186	2	-0.329	3
18-month	-0.314	1	-0.362	2	-0.441	3

(b) first difference

Series	ADF t-statistic	# of lags	ADF t-statistic	# of lags	ADF t-statistic	# of lags
WTI spot	-24.20^{***}	1	-19.07^{***}	2	-15.34^{***}	3
6-month	-23.26^{***}	1	-17.83^{***}	2	-14.99^{***}	3
12-month	-22.98^{***}	1	-17.80^{***}	2	-15.14^{***}	3
18-month	-19.28^{***}	1	-15.26^{***}	2	-13.02^{***}	3

Note: (i) The above tests are performed using log-levels of the prices; (ii) *, ** and ***denote that the null of a unit root is rejected at the 10%, 5% and 1% significance levels, respectively.

Table 2: **Estimated Unobserved Model of Spot and Futures prices**

parameters	model estimates
log likelihood	2891.87
Akaike (AIC) criterion	-4.29
Bayesian (BIC) criterion	-4.23
ρ_1	$1.2048(0.0264)^{***}$
ρ_2	$-0.2080(0.0269)^{***}$
ρ_{rp}	$0.9742(0.0060)^{***}$
σ^2_τ	$0.9597(0.3611)^{***}$
σ^2_c	$1.6829(0.6000)^{***}$
σ^2_{rp}	$0.0483(0.0037)^{***}$
σ^2_p	$0.5520(0.0296)^{***}$
σ^2_{f6}	$0.0042(0.0009)^{***}$
σ^2_{f12}	$0.0000(5.5E+10)$
σ^2_{f18}	$0.0064(0.0004)^{***}$
$\mu_{rp_{f6}}$	$0.0003(0.5331)$
$\mu_{rp_{f12}}$	$-0.2349(1.0425)$
$\mu_{rp_{f18}}$	$-0.4359(1.4239)$
β_{12}[a]	$1.3786(0.0190)^{***}$
β_{18}	$1.4020(0.0382)^{***}$
$cov_{\tau,c}$	$-0.9631(0.4818)^{**}$

Note: (i) Standard errors are in parentheses; (ii) *, ** and ***denote that the point estimate is significant at the 90%, 95% and 99% confidence levels, respectively.

[a]The loading factor for 6-month futures prices is normalized to be 1.

Table 3: **Out-of-sample Forecast Performance: 1994 - 2014**

(a) Out-of-Sample Forecast MSE

forecasting horizon	rolling [a]	forward recursive [b]	futures prices [c]	random walk
4-week	5.6754^{*2}	5.7471	5.3968	5.8447
8-week	12.6354	12.6434	12.3574	13.0592
12-week	21.5991	21.5326	21.4724	22.2394
16-week	30.8979	30.6886	30.9778	31.6885
20-week	40.1604	39.7054	40.4026	41.0125
24-week	48.8360	48.0105	49.1441	49.5805
28-week	55.7641	54.4636	55.9356	56.6205
32-week	60.4069	58.7147	60.4375	61.5599
36-week	63.6467	61.4869	63.3219	65.1098
40-week	65.7654	63.1187	65.0160	68.2042
44-week	67.0145	64.1373	66.1304	70.8562
48-week	68.0901^{*1}	65.1617	67.2564	74.2680

(b) Out-of-Sample Forecast ME

forecasting horizon	rolling[a]	forward recursive[b]	futures prices[d]	random walk
4-week	-0.0599	-0.0262	-0.0773	-0.1261
8-week	-0.1331	-0.1380	-0.1341	-0.2470
12-week	-0.2060	-0.2506	-0.2307	-0.3666
16-week	-0.2756	-0.3605	-0.3521	-0.4820
20-week	-0.3504	-0.4759	-0.4915	-0.6016
24-week	-0.4280	-0.5942	-0.6420	-0.7230
28-week	-0.5028	-0.7099	-0.7967	-0.8406
32-week	-0.5684^{*2}	-0.8164	-0.9440	-0.9483
36-week	-0.6290^{*2}	-0.9177	-1.0861	-1.0499
40-week	-0.6758^{**2}	-1.0050	-1.2131	-1.1370
44-week	-0.7101^{**2}	-1.0794	-1.3266	-1.2105
48-week	-0.7166^{**2}	-1.1258	-1.4120	-1.2555

Continued on the next page

Note: (i) *, ** and ***denote that the null hypothesis of equal predictive accuracy of compared models in Giacomini and White (2006) finite-sample unconditional test can be rejected at the 90%, 95% and 99% significance levels, respectively. 1 represents when compared to the random walk forecasts, 2 represents when compared to the futures price forecasts. HAC estimators of the variance for the test statistics are computed with Barlett kernel and a bandwidth of 100.

[a]Forecasts are generated from rolling estimation of the model, using 260 weeks (approximately 5 years) of data. The first estimation uses data from March 29, 1989 to March 18, 1994. The estimation sample period is rolled forward at weekly frequency until Jan 31, 2014.

[b]Forecasts are generated from forward recursive estimation of the model. The first estimation uses data from March 29, 1989 to March 18, 1994. The estimation sample period is extended at weekly frequency until Jan 31, 2014.

[c]The futures prices used for forecasting at 1- to 12-month futures prices, quoted at weekly frequency.

Table 3: **Out-of-sample Forecast Performance: 1994 - 2014 - Continued**

(c) Out-of-Sample Forecast MAE

forecasting horizon	rolling[a]	forward recursive[b]	futures prices[d]	random walk
4-week	1.6888^{***2}	1.6839	1.6190	1.6847
8-week	2.3573	2.3605	2.3333	2.3849
12-week	2.9051	2.9244	2.9127	2.9932
16-week	3.3313	3.3676	3.3804	3.4448
20-week	3.6822	3.7544	3.7935	3.8514
24-week	4.0006^{*1}	4.1160	4.1728	4.2328
28-week	4.3275^{**1,*2}	4.4394	4.5327	4.5949
32-week	4.5295^{**1}	4.6386	4.7561	4.8131
36-week	4.6505^{**1}	4.7754	4.9069	4.9671
40-week	4.7870^{***1}	4.9123	5.0766	5.1889
44-week	4.9441^{***1}	5.1084	5.2744	5.4196
48-week	5.1367^{***1}	5.3091	5.4649	5.7027

(d) Fraction of Smaller Absolute Forecasting Errors

	benchmark: futures prices[d]		benchmark: random walk	
forecasting horizon	rolling[a]	forward recursive[b]	rolling[a]	forward recursive[b]
4-week	0.4489	0.4576	0.4990	0.4817
8-week	0.4624	0.4624	0.5000	0.4884
12-week	0.4971	0.4875	0.5222	0.4798
16-week	0.5337	0.5376	0.5164	0.4692
20-week	0.5713	0.5665	0.5299	0.4846
24-week	0.5906	0.5848	0.5482	0.4817
28-week	0.5934	0.6127	0.5520	0.4942
32-week	0.5963	0.6320	0.5366	0.4913
36-week	0.6021	0.6445	0.5395	0.4855
40-week	0.6224	0.6580	0.5645	0.4904
44-week	0.6252	0.6532	0.5617	0.4904
48-week	0.6175	0.6647	0.5848	0.5010

Table 4: **Out-of-sample Forecast Performance Robustness: 1994 - 2014**

(a) Out-of-Sample Forecast MSE

forecasting horizon	rolling sample 1[ab]	rolling sample 2[ac]	futures prices[d]	random walk
4-week	5.8064	5.3114^{*1}	5.3968	5.8447
8-week	13.4625	12.2535^{*1}	12.3574	13.0592
12-week	22.9402	21.3038	21.4724	22.2394
16-week	32.5562	30.6874	30.9778	31.6885
20-week	41.8391	39.8763	40.4026	41.0125
24-week	50.0224	48.2978^{*2}	49.1441	49.5805
28-week	56.2174	54.7817^{**2}	55.9356	56.6205
32-week	60.4301	59.1245^{***2}	60.4375	61.5599
36-week	63.1276	61.9042^{***2}	63.3219	65.1098
40-week	64.9418	63.5531^{***2}	65.016	68.2042
44-week	66.033	64.6714^{***2}	66.1304	70.8562
48-week	67.368	65.8049^{**2}	67.2564	74.268

(b) Out-of-Sample Forecast ME

forecasting horizon	rolling sample 1[ab]	rolling sample 2[ac]	futures prices[d]	random walk
4-week	-0.1577	-0.0939	-0.0773	-0.1261
8-week	-0.2673	-0.1886	-0.1341	-0.247
12-week	-0.3773	-0.2856	-0.2307	-0.3666
16-week	-0.4846	-0.3814	-0.3521	-0.482
20-week	-0.5973	-0.484	-0.4915	-0.6016
24-week	-0.7128	-0.5907	-0.642	-0.723
28-week	-0.8256	-0.6956	-0.7967	-0.8406
32-week	-0.9291	-0.7921^{*2}	-0.944	-0.9483
36-week	-1.0272	-0.8842^{***2}	-1.0861	-1.0499
40-week	-1.1113	-0.9629^{***2}	-1.2131	-1.137
44-week	-1.1823	-1.0292^{***2}	-1.3266	-1.2105
48-week	-1.2251^{**2}	-1.0679^{***2}	-1.412	-1.2555

Continued on the next page

Note: (i) *, ** and ***denote that the null hypothesis of equal predictive accuracy of compared models in Giacomini and White (2006) finite-sample unconditional test can be rejected at the 90%, 95% and 99% significance levels, respectively. 1 represents when compared to the random walk forecasts, 2 represents when compared to the futures price forecasts. HAC estimators of the variance for the test statistics are computed with Barlett kernel and a bandwidth of 100.

[a]Forecasts are generated from rolling estimation of the model, using 260 weeks (approximately 5 years) of data. The first estimation uses data from March 29, 1989 to March 18, 1994. The estimation sample period is rolled forward at weekly frequency until Jan 31, 2014.

[b]Sample 1 includes the spot, 3-, 4- and 6-month futures prices

[c]Sample 2 includes the spot, 3-, 6- and 12-month futures prices

[d]The futures prices used for forecasting at 1- to 12-month futures prices, quoted at weekly frequency.

Table 4: **Out-of-sample Forecast Performance Robustness: 1994 - 2014 - Continued**

(c) Out-of-Sample Forecast MAE

forecasting horizon	rolling sample 1[ab]	rolling sample 2[ac]	futures prices[d]	random walk
4-week	1.67	1.6206^{**1}	1.619	1.6847
8-week	2.4102^{*2}	2.3233	2.3333	2.3849
12-week	2.9899	2.8963	2.9127	2.9932
16-week	3.4702	3.366	3.3804	3.4448
20-week	3.8945	3.7603	3.7935	3.8514
24-week	4.2811	4.1226	4.1728	4.2328
28-week	4.6004	4.4633	4.5327	4.5949
32-week	4.8007	4.6717^{**2}	4.7561	4.8131
36-week	4.9362	4.8106^{**2}	4.9069	4.9671
40-week	5.098	4.96^{***2}	5.0766	5.1889
44-week	5.2765	5.1526^{***2}	5.2744	5.4196
48-week	5.4825	5.3498^{**2}	5.4649	5.7027

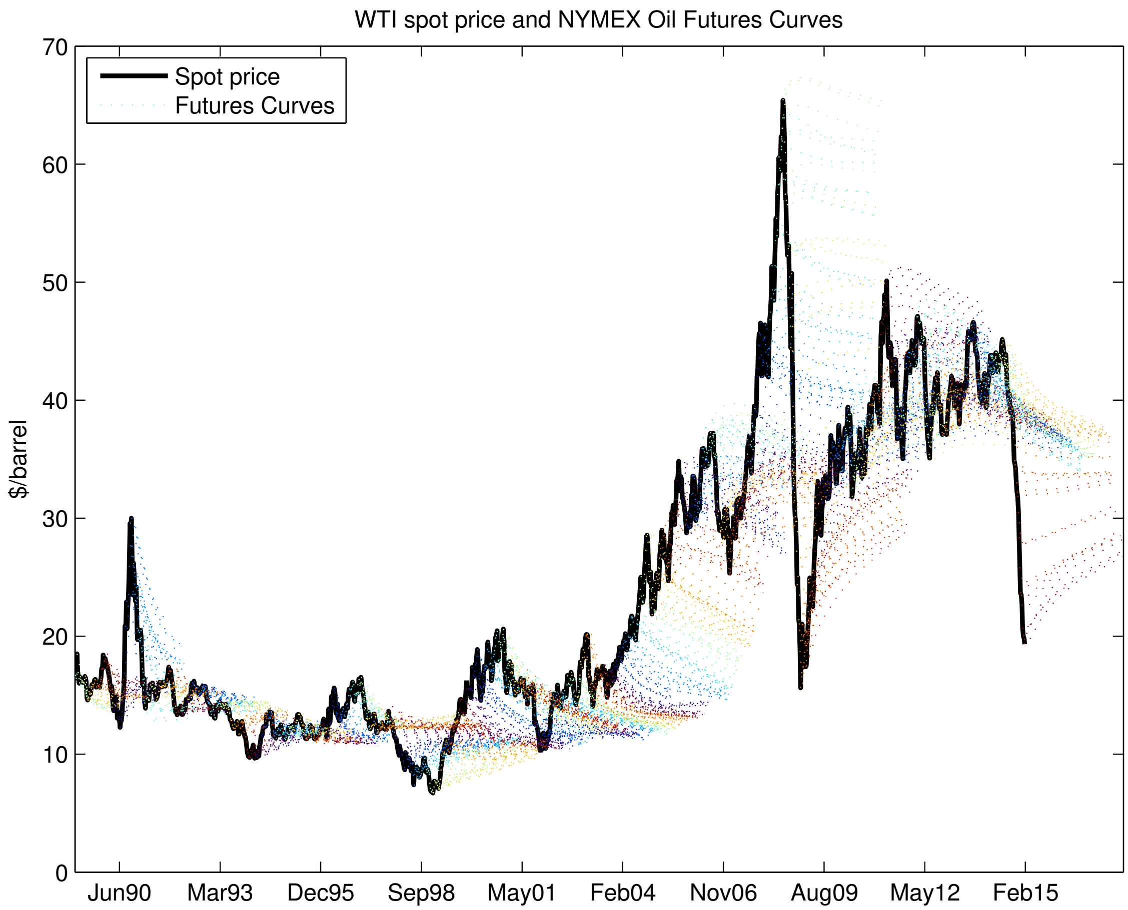
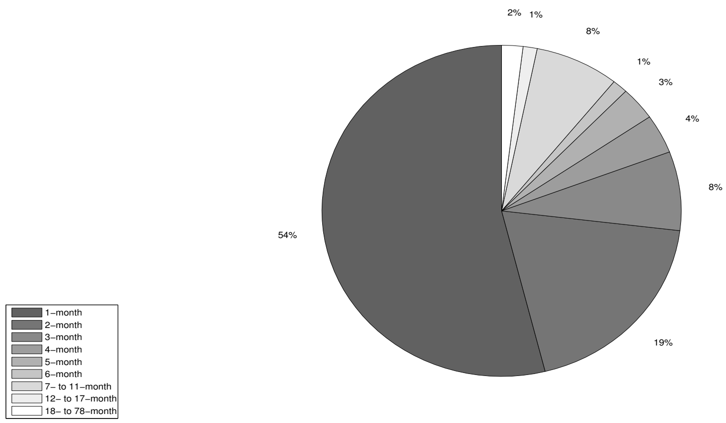
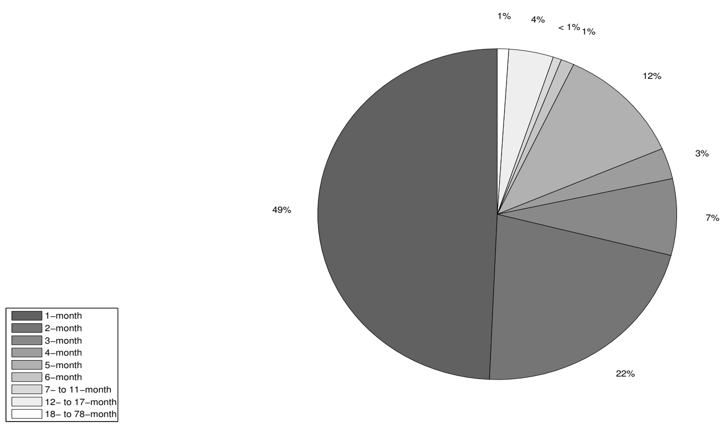


Figure 1: WTI spot price with futures curves

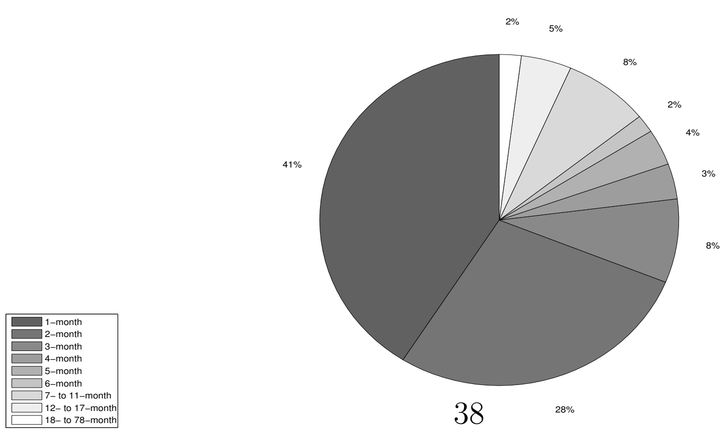
Source: Datastream. Weekly real prices (calculated by the author).



(a) March 27 1992



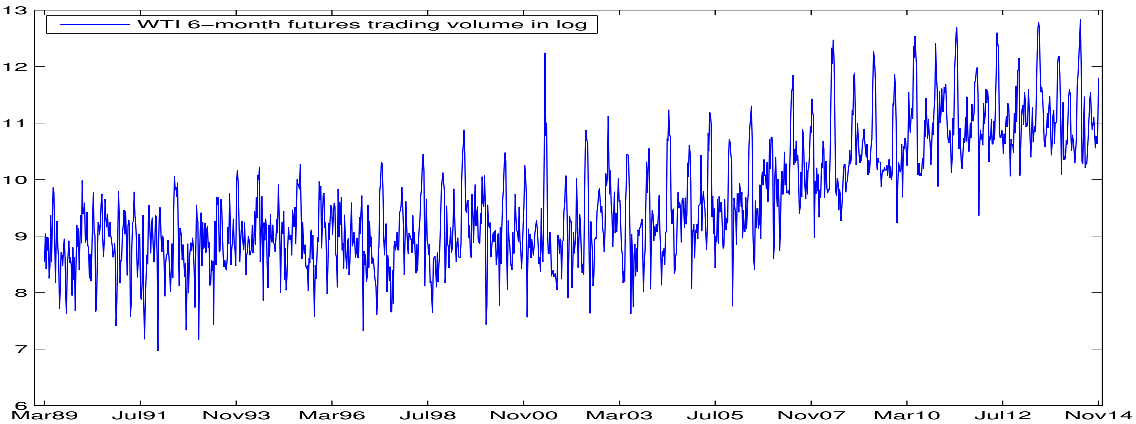
(b) June 21 2001



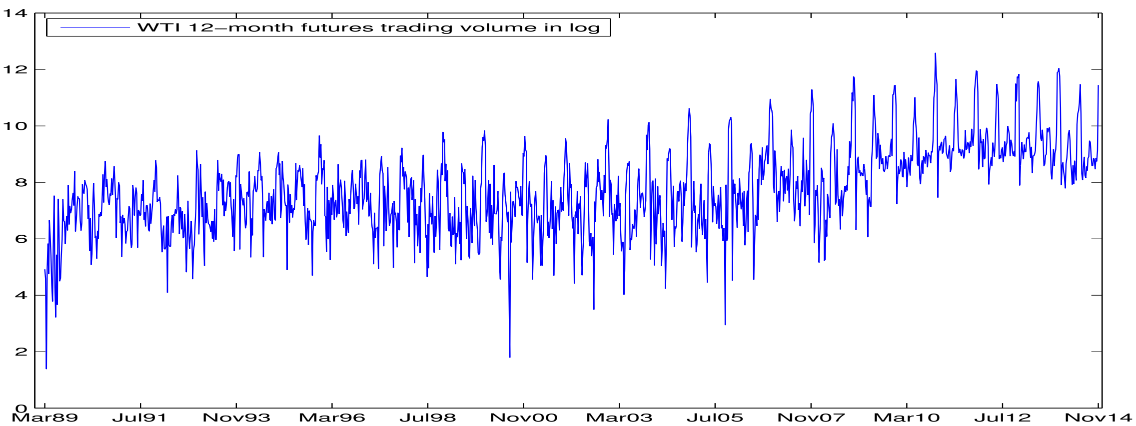
(c) October 14 2011

Figure 2: WTI Futures Contracts Trading Distribution on a Typical Day

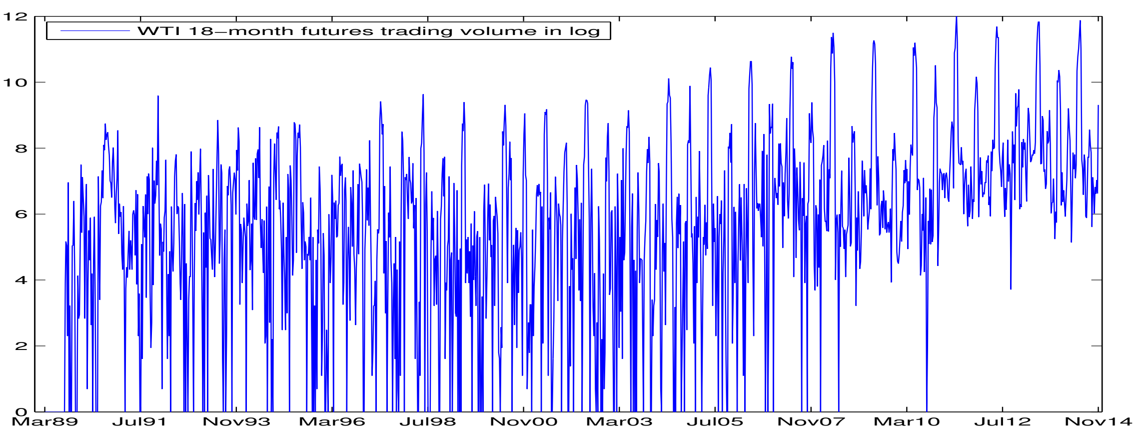
Note: The trading volume of each category as a share of the total trading volume on a typical day.
Source: Datastream. Daily trading volume.



(a) 6-month



(b) 12-month



(c) 18-month

Figure 3: WTI Futures Contracts Weekly Trading Volume

Note: When the trading volume is NA or zero, the log is plotted in the figures as zero.
Source: Datastream. Weekly trading volume.

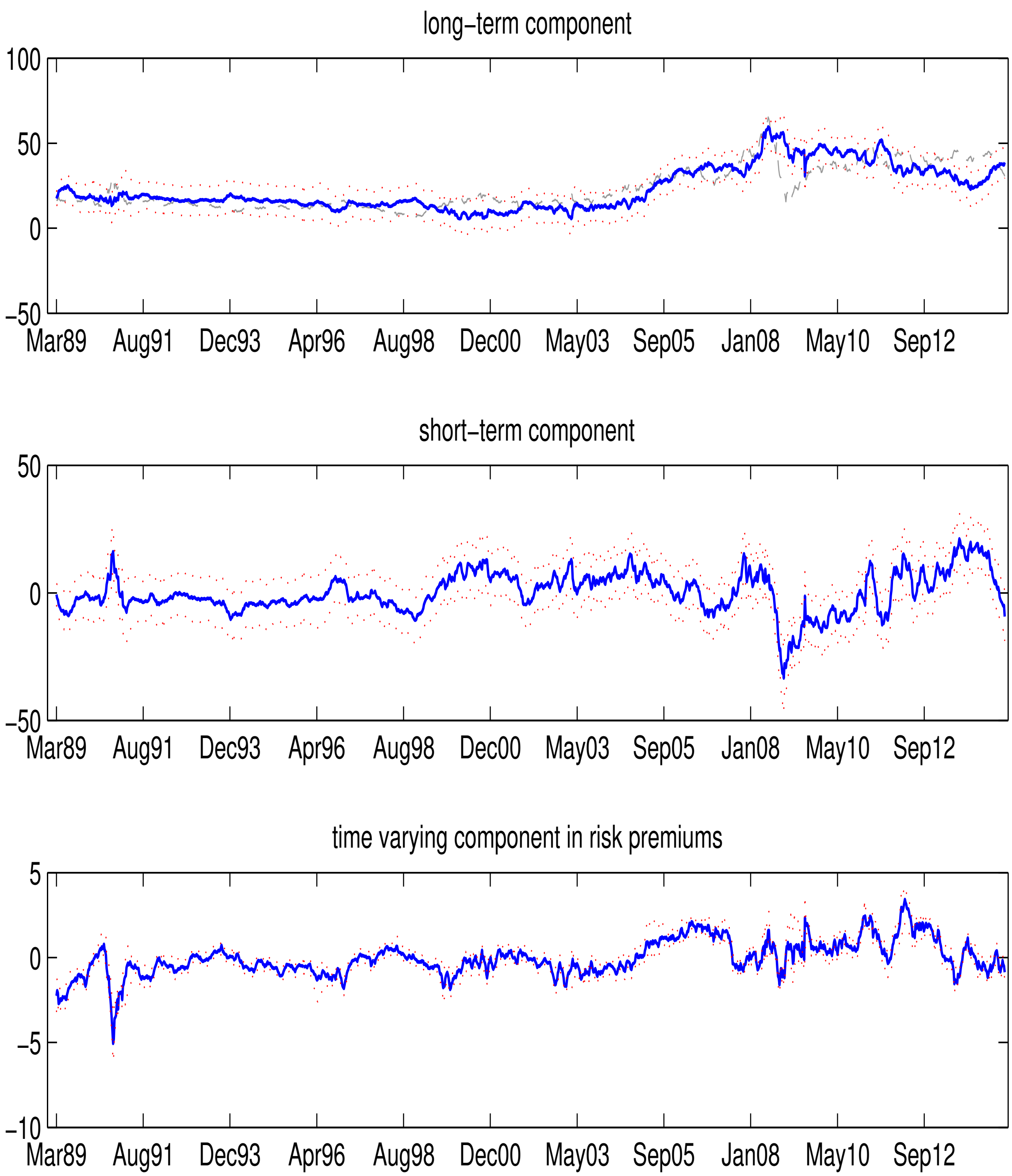


Figure 4: Estimated Unobserved Components

Note: The grey line plotted along with the long-term component is WTI spot price.

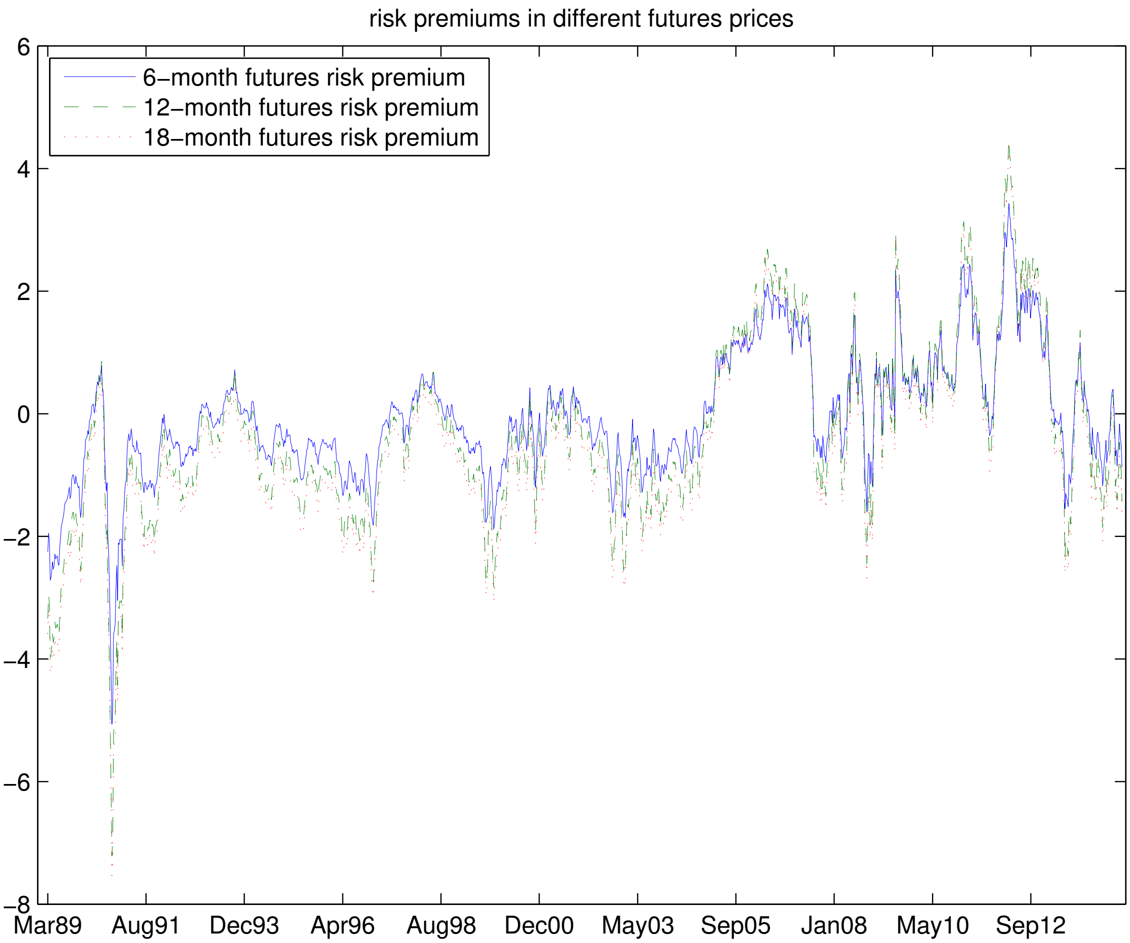


Figure 5: Estimated Risk Premiums of Different Futures Prices

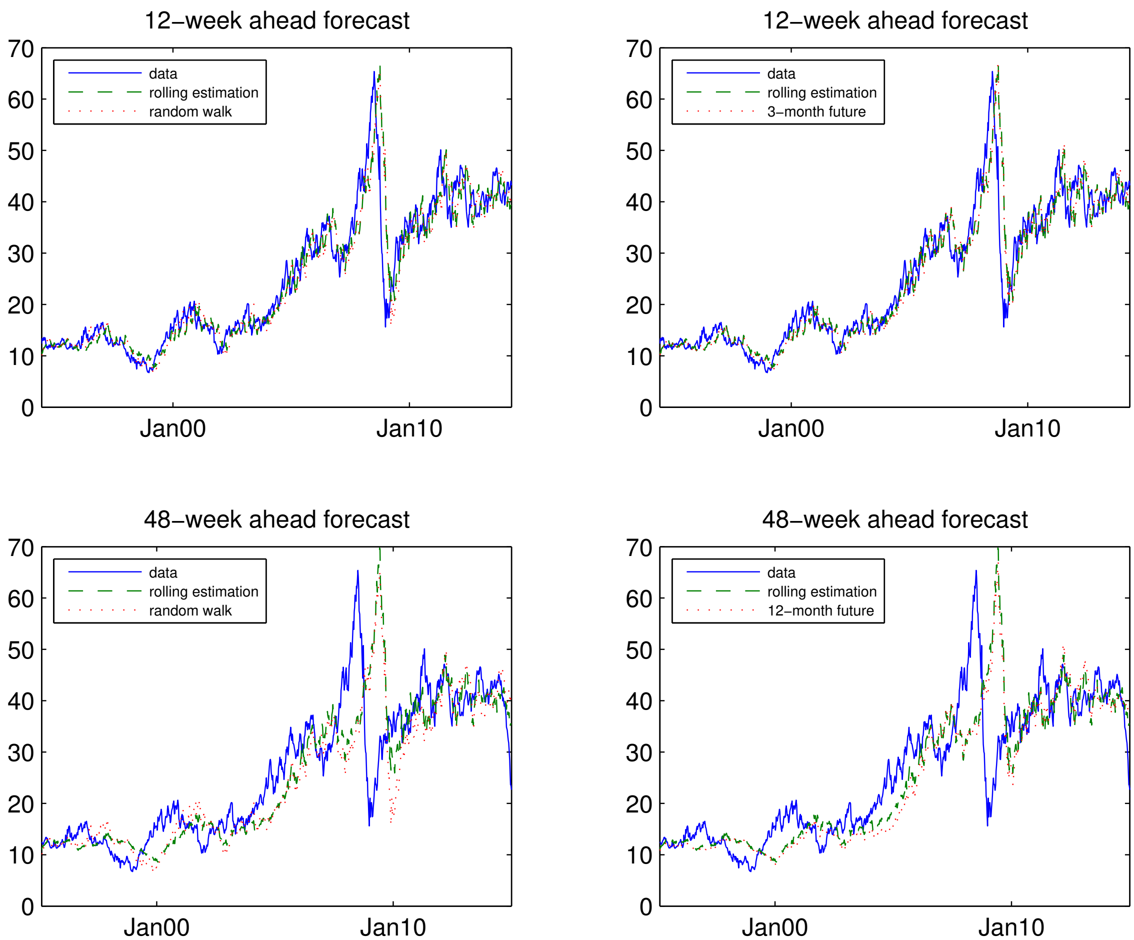


Figure 6: Comparison of Different Forecasts at Selected Forecasting Horizons

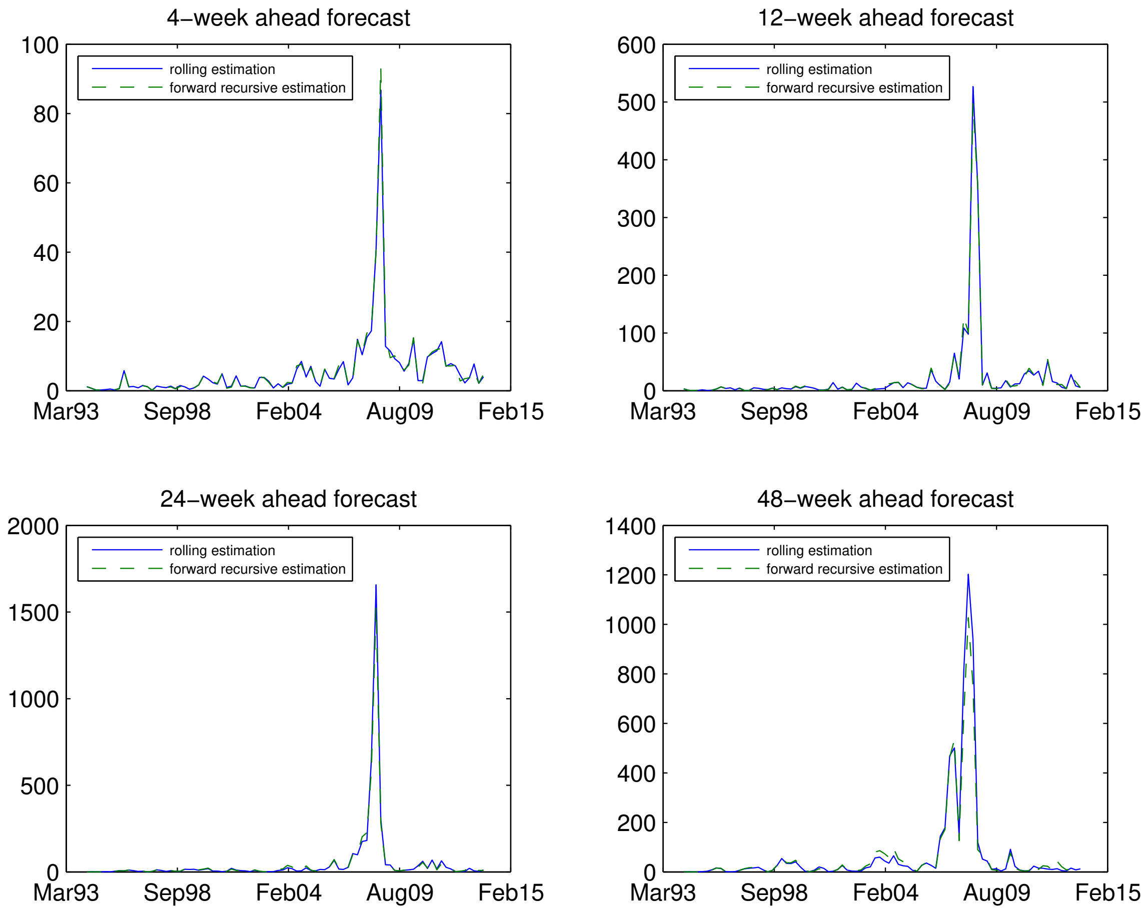


Figure 7: 12-week MSE at Different Horizons using Model Forecasts

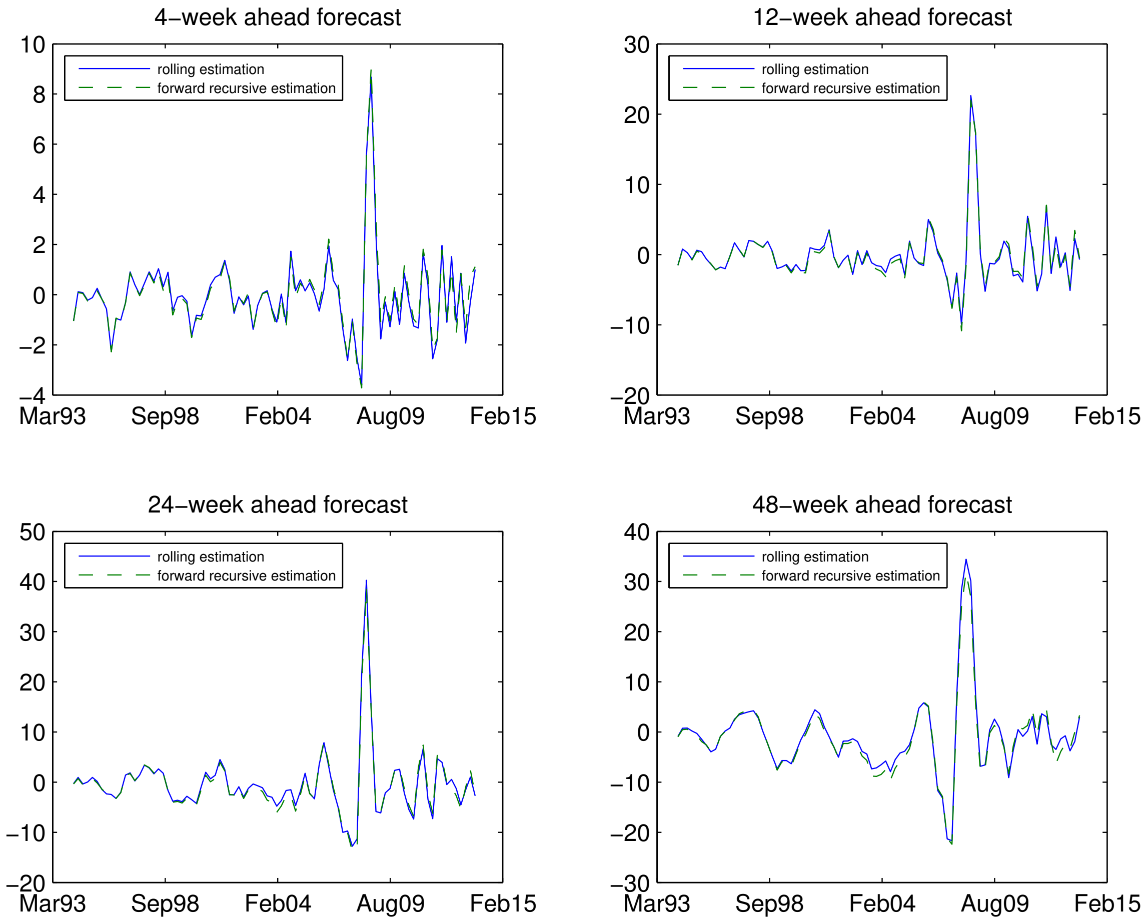


Figure 8: 12-week ME at Different Horizons using Model Forecasts

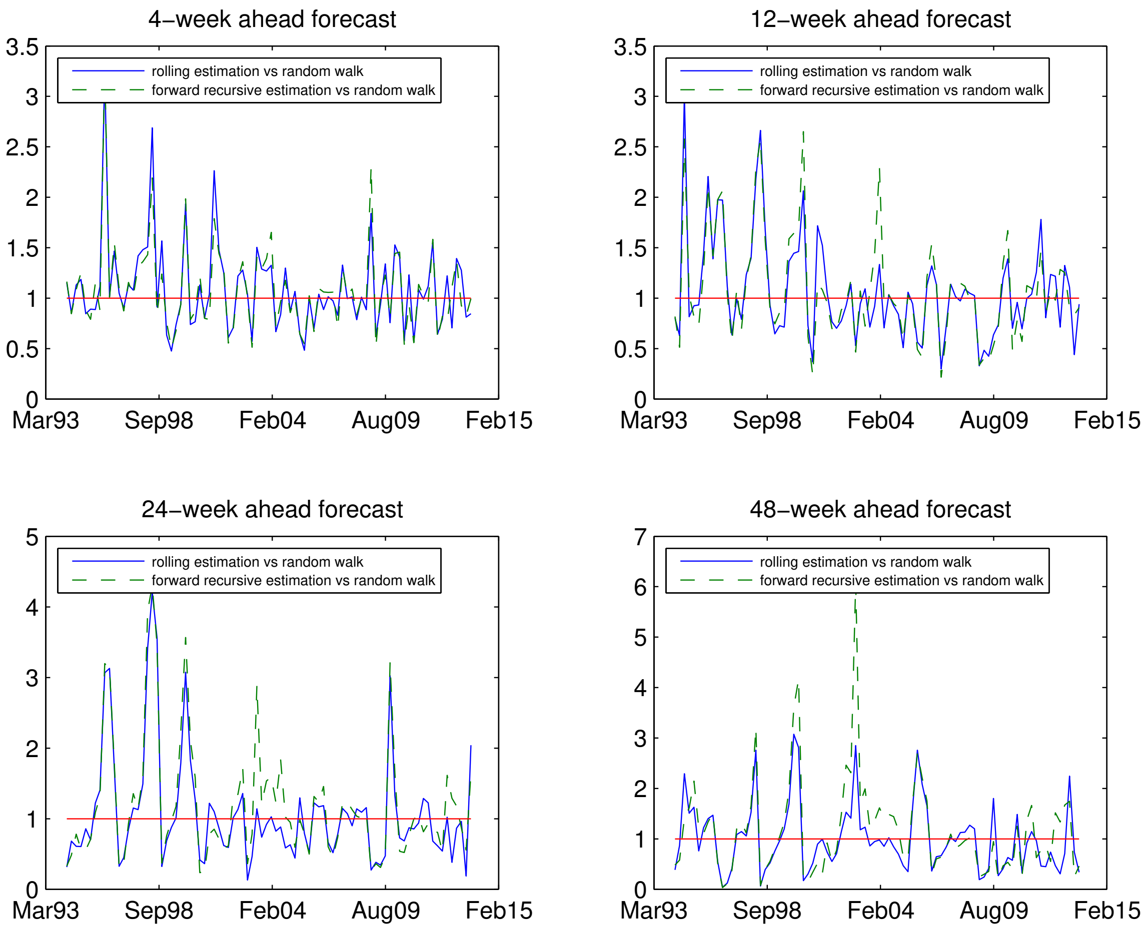


Figure 9: The model relative to the random walk: 12-week MSE ratios at Different Horizons

Note: Model forecasts 12-week MSE to random walk counterpart ratio is plotted with the red line =1.

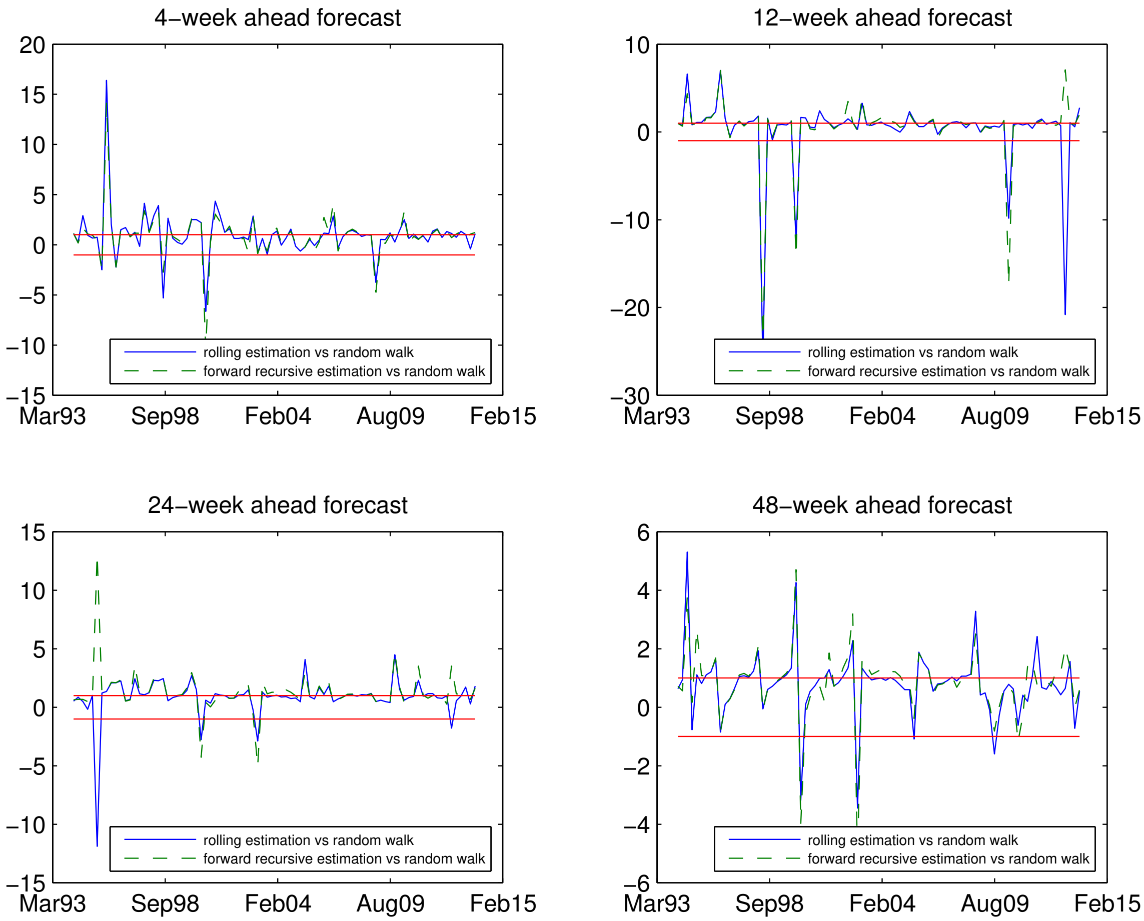


Figure 10: Model relative to random walk: 12-week ME ratios at Different Horizons using Model Forecasts

Note: The model forecasts 12-week ME to the random walk counterpart ratio is plotted with the red lines =1 and -1.

References

Alquist, Ron and Lutz Kilian (2010), "What do we learn from the price of crude oil futures?" *Journal of Applied Econometrics*, 25, 539–573.

Alquist, Ron, Lutz Kilian, and Robert J. Vigfusson (2013), "Forecasting the price of oil." In *Handbook of Economics Forecasting* (Graham Elliott and Allan Timmermann, eds.), volume 2, 427–507, North-Holland.

Bopp, Anthony E. and George M. Lady (1991), "A comparison of petroleum futures versus spot prices as predictors of prices in the future." *Energy Economics*, 13, 274–282.

Bowman, Chakriya and Aasim M. Husain (2004), "Forecasting commodity prices: Futures versus judgment." IMF Working Paper.

Brennan, M.J. (1958), "The supply of storage." *The American Economic Review*, 48, 50–72.

Caporale, Guglielmo, Davide Ciferri, and Alessandro Girardi (2014), "Time-Varing Spot and Futures Oil Price Dynamics." *Scottish Journal of Political Economy*, 61, 28.

Carlson, Murray, Zeigham Khokher, and Sheridan Titman (2007), "Equilibrium exhaustible resource price dynamics." *The Journal of Finance*, LXII, 1663–1703.

Chernenko, Sergey V., Krista B. Schwarz, and Jonathan H. Wright (2004), "The information content of forward and futures prices: Market expectations and the price of risk." International Finance Discussion Papers No.808.

Chinn, Menzie D. and Olivier Coibion (2014), "The predictive content of commodity futures." *The Journal of Futures Market*, 34, 607–636.

Coppola, Andrea (2008), "Forecasting oil price movements: Exploiting the information in the futures market." *The Journal of Futures Markets*, 28, 34–56.

Cortazar, Gonzalo and Eduardo S. Schwartz (2003), "Implementing a stochastic model for oil futures prices." *Energy Economics*, 25, 215–238.

Durbin, James and Siem Jan Koopman (2012), *Time Series Analysis by State Space Methods*, second edition. Oxford : Oxford University Press.

Fama, Eugene F. and Kenneth R. French (1987), "Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage." *The Journal of Business*, 60, 55–73.

Fama, Eugene F. and Kenneth R. French (1988), "Permanent and temporary components of stock prices." *Journal of Political Economy*, 96, 246–273.

Figuerola-Ferretti, Isabel and Jesus Gonzalo (2010), "Modelling and measuring price discovery in commodity markets." *Journal of Econometrics*, 158, 95–107.

French, Kenneth R. (1986), "Detecting spot price forecasts in futures prices." *The Journal of Business*, 59, S39–S54.

Giacomini, Raffaella and Halbert White (2006), "Tests of Conditional Predictive Ability." *Econometrica*, 74, 1545–1578.

Giannone, Domenico, Lucrezia Reichlin, and David Small (2008), "Nowcasting: The real-time informational content of macroeconomic data." *Journal of Monetary Economics*, 55, 665–676.

Hamilton, James D. (1994), *Time Series Analysis*. Princeton, N.J. : Princeton University Press.

Hamilton, James D. and Jing Cynthia Wu (2014), "Risk premia in crude oil futures prices." *Journal of International Money and Finance*, 42, 9–37.

Harvey, Andrew C (1989), *Forecasting, structural time series models, and the Kalman filter*. Cambridge University Press.

Harvey, Andrew C and R G Pierse (1984), "Estimating Missing Observations in Economic Time Series." *Journal of the American Statistical Association*, 79, 125–131.

Herce, Miguel, John E. Parsons, and Robert C. Ready (2006), "Using futures prices to filter short-term volatility and recover a latent, long-term price series for oil." MIT Center for Energy and Environmental Policy Research Working Paper 06-005.

Keynes, John Maynard (1930), *A Treatise on Money*. Macmillan, Lodon.

Malkiel, Burton G. (2003), "The efficient market hypothesis and its critics." *Journal of Economic Perspectives*, 17, 59–82.

Moosa, Imad and Nabeel Al-Loughani (1994), “Unbiasedness and time varying risk premia in the crude oil futures market.” *Energy Economics*, 16, 99–105.

Peck, A.E. (1985), “Economic role of traditional commodity futures market.” In *Futures Markets: Their Economic Role* (A.E. Peck, ed.), chapter 1, 1–81, American Enterprise Institute for Public Policy Research, Washington, D.C.

Pindyck, Robert S. (1999), “The long-run evolution of energy prices.” *Energy Journal*, 20, 1–27.

Pindyck, Robert S. (2001), “The dynamics of commodity spot and futures markets: A primer.” *The Energy Journal*, 22, 1–29.

Reeve, Trevor A and Robert J. Vigfusson (2011), “Evaluating the forecasting performance of commodity futures prices.” International Finance Discussion Papers No.1025.

Reichsfeld, David A. and Shaun K. Roache (2009), “Do commodity futures help forecast spot prices?” IMF Working Paper.

Schwartz, Eduardo S (1997), “The stochastic behavior of commodity prices: Implications for valuation and hedging.” *Journal of Finance*, 52, 923–73.

Silvério, Renan and Alexandre Szklo (2012), “The effect of the financial sector on the evolution of oil prices: Analysis of the contribution of the futures market to the price discovery process in the WTI spot market.” *Energy Economics*, 34, 1799–1808.

Stevens, Jason (2013), "The benefits of storage and non-renewable resource price dynamics." *Canadian Journal of Economics*, 46, 239–265.

Tomex, William G. and Roger W. Gray (1970), "Temporal relationships among prices on commodity futures markets: Their allocative and stabilizing roles." *American Journal of Agricultural Economics*, 52, 372–380.

Working, Holbrook (1942), "Quotations on commodity futures as price forecasts." *Econometrica*, 10, 39–52.